

정렬반안 망을 이용한 성능이 향상된 스위치설계

정희원 권승탁*

Design of Speed-up Switch Using Sort Banyan Networks

Seung-Tag Kwon* Regular Member

요 약

통신망은 스위칭 장치들을 상호 연결하여 구성한다. 스위칭 장치는 경로배정 정보에 의하여 입력포트와 출력포트 사이를 연결하는 역할을 한다. 그런데 반안 망을 사용하는 대부분의 스위치 망은 두개의 셀 들이 같은 길을 사용하려고 시도할 때 내부 충돌이 발생한다.

본 논문은 입력과 출력사이에 목적지가 같은 두 개의 셀 들이 충돌이 없도록 동시에 두 개 경로를 설정하여 셀 처리율을 향상시키는 개선된 정렬 반안 망을 제안하고 설계하였다. 하드웨어 설계는 2개의 4×4 정렬 부 블록들과 1개의 8×8 스위치 망 블록으로 구성하였는데 기존의 정렬망인 배치 반안 망보다 스위치의 셀 처리능력은 4% 향상되었고 정렬 부의 하드웨어 복잡도는 반으로 감소하였다.

키워드: 스위치 설계

ABSTRACT

A network is made up of interconnected switching units. The role of a switching unit is to set up a connection between an input port and an output, according to the routing information. But then the most switching network use Banyan switch, their occurs the internal blocking, which attempts to use the same link two cells.

This paper proposed and designed for a improvement Batch-Banyan network which can routed two path assignment between its input ports and output ports without only blocking. The network is constructed of two sorting blocks (4×4), one switch network(8×8)block. As a result, the switch network performance increased 4% reduced to half of the hardware complexity of sorting boxes when compare the new switching system with Batcher-Banyan network system.

1. 서 론

스위치는 대규모 통신망이나 병렬 컴퓨터를 설계하는데 중요한 역할을 담당한다. 스위치 망은 경로배정(routing) 알고리즘에 의하여 스위치를 배치한 다음 서로의 노드를 연결하여 구성한다^{[1][2]}. 스위치 망들 중에서 반안 망은 완전 셔플 망으로서 알고리즘의 변화 없이 소형에서 대형 망까지 사용할 수 있으며 알고리즘이 간단하여 하드웨어 구현이 용이하다.

그러나 출력주소가 같은 여러 개의 셀 들이 스위치 망을 통해 이동할 때 셀들 간에 충돌(blocking)을 일으킨다^[3]. 따라서 목적지가 같은 모든 셀 중 한 개

를 제외한 셀들은 블록화 하여 제거(drop)시키거나 임시버퍼에 저장하여 처리하여 셀들 간에 충돌을 방지하는 방법이 필요하다. 이러한 충돌을 방지하기 위한 방법으로는 내부 속도를 증가시키거나, 입출력 및 내부에 버퍼를 두거나, 스위치 앞단에 정렬(sorted) 망을 두어 비 충돌(non-blocking) 망을 구성하는 방법들이 주로 쓰인다^[4].

그러나 반안 망이 비 충돌 망일지라도 출력주소가 같은 셀 들이 두 개 이상 동시에 통과할 수 없는 구조를 갖기 때문에 입력링크에 무작위로 도착한 셀 들이 출력에 도착할 수 있는 최대 처리율은 58.6%로 제한된다^[5].

* 서남대학교 컴퓨터 정보통신학과 조교수 정보통신 시스템 연구실(stkwon@tiger.seonam.ac.kr)
논문번호 : 020470-1029, 접수일자 : 2002년 10월 29일

이러한 현상은 입력 단에 하나의 선입선출(FIFO: First In First Out)큐를 두고 입출력 포트 $N \times N$ 구조를 사용하는 경우에 발생하므로 이 조건을 완화시킴으로써 성능 저하요인을 감소 또는 제거시키려는 연구가 진행되어져 왔는데 이러한 방법으로는 입력 원활화 (Input smooth) 방식^{[6][7]}과 목적지별 큐잉 방식 등이 연구되어져 왔다.

본 논문은 반안 망의 내부 구조를 변경하여 스위치의 성능을 향상시키는 방법을 제안하고 하드웨어로 설계한다.

기존의 대표적인 정렬 망은 베처반안 망으로, 이 망은 셀 들이 반안 망에 입력되기 전 셀 들의 출력주소를 순차적으로 배열하여 충돌을 방지 할 수 있지만 입력 링크에 도착한 셀 중에서 출력주소가 같은 셀들 중 한 개의 셀만 처리하기 때문에 셀 처리율 향상에 큰 도움이 되지 못한다^[8]. 따라서 본 논문은 기존의 정렬망인 베처반안 망의 구조를 개선하여 동시에 두 개의 셀들을 처리하는 방법을 제안하여 베처반안 망에 비하여 셀 처리율을 향상시키고 하드웨어 복잡도를 줄일 수 있는 방법을 제안하고 하드웨어로 설계한다.

제안된 정렬반안 망 스위치는 MyCAD를 사용하여 설계 및 검증을 한다.

본 논문의 구성은 2장에서 베처반안 스위치의 구조와 동작에 관하여 알아보았고 3장에서는 제안한 정렬반안 망의 동작과 논리회로 시뮬레이션 결과를 살펴보고, 4장에서는 스위치 망의 성능을 평가하였다. 그리고 5장에서는 결론을 맺었다.

II. 베처 반안 스위치의 구조와 동작

베처반안 망은 반안 망의 비 충돌 조건인 “모든 입력 셀들이 인접포트이고 출력주소가 크기 순에 따라 증가하거나 감소할 때 서로 다른 경로를 이용하여 이동한다.”는 정의를 만족시키기 위하여 정렬 망을 반안 망 앞단에 배치한 망이다. 이 망은 입력링크에 도착한 셀들이 망에 입력되기 전에 입력된 셀들의 출력주소를 순차적으로 배열하여 셀들이 반안 망 내부에서 충돌이 발생하지 않도록 구성된 스위치 망이다. 본 절에서는 반안 망과 정렬 망의 구조와 동작을 설명한다.

2.1 반안 망의 구조

반안 망은 완전 셔플 망으로 연결특성을 살펴보면 다음과 같다. 반안 망의 첫 번째 단은 $N \times N$ (N :입출

력 링크 수)블록을 포함하고 두 번째 단은 $(N/2) \times (N/2)$ 두 개의 하위블록으로 이루어진다. 망을 구성하는 과정은 크기가 2×2 인 하위블록이 될 때까지 재귀적으로 적용된다. 따라서 망 전체적인 연결은 순열구조를 갖추고 있다.

반안 망의 입력 X_k 이고 Y_k 일때 $Y_{n,d}$ 는 다음과 같이 정의된다.

$$Y_{d,d} = X_{d,d} \tag{1}$$

$$Y_{n,d} = \tau_{n/d} (X_{d,d} \times Y_{n/d,d})$$

예를 들어 각각 16개의 입출력포트와 2×2 로 구성된 반안 망은 정의 식에 따라 구성하면 식 $Y_{16,2}$ 와 같이 표현하며 $Y_{16,2}$ 의 내부에는 $Y_{8,2}$ 가 있고, 또 다시 $Y_{8,2}$ 는 $Y_{4,2}$ 로 구성한다.

반안 망의 단 수는 $K = \log_2 N$ 개이다. 여기서 b 는 스위칭 소자의 크기이고 N 은 스위치의 크기이다. 그러므로 단위 교환소자가 2×2 스위치로 $N \times N$ 의 교환 망을 구성하면 스위치의 단은 $\log_2 N$ 단으로 이루어진다.

정의된 식에 의하면 반안 망은 단일 경로를 갖고 있으므로 여러 개의 셀 들이 동시에 전달될 경우 망 내부에서 충돌이 발생한다.

따라서 충돌을 방지하기 위한 방법이 필요하다. 반안 망은 다음 조건을 만족할 때 비 충돌 특성을 갖는다.

반안 망 스위치의 또 다른 특징은 자기경로배정 (self-routing)을 사용 할 수 있다^[9]. 자기경로배정이란 패킷의 선두부분의 출력주소 h_1, h_2, \dots, h_k 에 의하여 출력포트 ot_1, ot_2, \dots, ot_k 에 전달하는 방법이다. 경로배정을 위하여 자기경로배정 주소 b_k 번째가 출력 n_k 를 지정 받기 위하여 노드 h_k 번째에서 사용되었다고 가정하면, 각 스위치에서 그에 맞는 비트를 비교함으로써 상호접속 망에서 단일경로를 따라 경로를 배정한다.

각 단의 노드는 자기경로 배정주소 b_1, b_2, \dots, b_k 를 자신의 단에 맞도록 비트를 추출한다. 자기경로 배정에서 단일 스위칭 소자 $X_{d,d}$ 의 동작을 살펴보면 목적지 주소 비트에 의한 자기경로배정에 의하여 i 번째 단에 입력되는 패킷의 목적지 주소 중 i 번째 비트 값에 따라 “0”일 경우 위쪽 출력포트로, “1”일 경우 아래쪽 출력포트로 스위칭 한다.

2.2 정렬 망

정렬 망¹⁸⁾은 스위치 망에 입력되는 셀들의 출력 주소에 따라 순차적으로 배열해주는 역할하며 셀들을 출력주소 순으로 정렬하여 스위치 망을 통과할 때 충돌이 발생하지 않도록 스위치 망 전단부에 설치한다.

그림 1은 정렬망인 베처반안 망의 구조를 보여주고 있다.

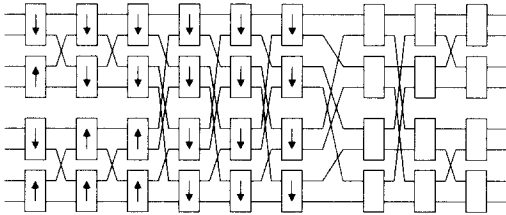


그림 1. 베처반안 망

베처반안 망의 구성은 8x8 스위치 일 경우 그림 1에서 보는 바와 같이 상향 정렬 소자 6개 하향 정렬 소자 18개로 합 24개의 정렬소자로 구성한다. 베처반안 망의 정렬 망의 단은 $\log_2 N \cdot (\log_2 N) / 2$ 의 크기를 갖는다.

III. 제안된 스위치 설계 및 시뮬레이션

기존의 정렬망인 베처반안 망은 반안 망의 비 충돌 특성을 만족시키기 위하여 반안 망 앞단에 정렬 망을 배치하여 설계하여 스위치 망을 통과 시켰으나 출력 주소가 같은 셀들은 한 개의 셀 밖에 통과 할 수 없으므로 스위치 망의 성능이 저하된다.

따라서 본 논문에서는 그림 2와 같이 하드웨어 양이 줄어들고 스위치의 성능이 향상된 스위치를 설계 하였다.

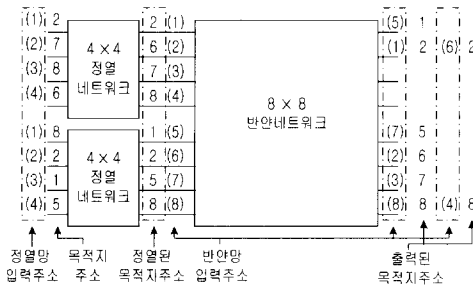


그림 2. 제안된 정렬 반안 스위치구조

제안된 스위치 망의 구성과 동작을 살펴보면 스위치 망 부분은 반안 망의 구조를 개선하여 같은 출

력을 가지는 두 개의 셀들이 동시에 통과하도록 구성하였고 이에 맞도록 정렬 부는 베처반안 망의 정렬 부의 구조를 개선하여 설계하였다.

제안된 스위치 망의 구조는 입출력 포트가 8x8개를 가지고 있으며 그림 3에 보는바와 같이 스위치 망의 전단 부에는 2 블록의 4x4의 정렬 망과 동시에 두개의 경로를 갖을 수 있는 8x8 스위치 망으로 구성하였다.

3.1 전체 회로의 동작

8x8 전체 정렬 반안 망에서 스위치 망은 $\log_2 N \cdot (\log_2 N - 1) / 2$ 단으로 구성되며 정렬 망에서는 제안된 정렬 반안 망 스위치의 동작은 그림 3과 같이 반안 망의 전단부 4x4의 정렬 망 2단에서 각각의 출력주소로 정렬한 다음 두개의 경로를 갖도록 설계된 8x8 반안 망을 통해 셀을 전달한다.

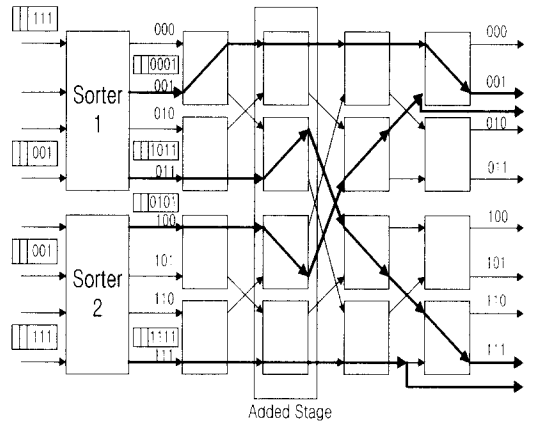


그림 3. 제안된 정렬반안 망의 동작

이 망의 구조는 4x4 정렬 망에서 정렬된 셀들이 출력포트에서 8x8 반안 망의 입력포트와 연결되어 있으며, 임의의 입력포트와 출력포트 사이에 2개의 경로를 가지고 있어 두개의 셀을 동시에 전송할 수 있다.

3.2 스위치 회로 설계

제안된 스위치는 입출력 포트의 수가 8x8개를 갖도록 설계하였다. 전체구조는 4x4 정렬 망 두 개와 두개의 셀을 동시에 처리할 수 있는 개선된 반안 망을 사용하였다.

1) 비교기 설계

8x8 베처반안 망에서 비교기 소자는 24개로 구성

되어 있으나, 제안된 스위치 망은 뒤에 연결되는 반안 망이 두 개의 경로를 갖기 때문에 12개의 비교기만을 사용하여 구성할 수 있다. 따라서 배저반안 망보다 하드웨어 양이 1/2로 줄었음을 볼 수 있다.

4x4 정렬 망의 동작을 살펴보면 다음과 같다. 입력포트에 목적지 주소를 가진 셀이 입력되면, 정렬 네트워크 안에서 주소가 2진 정렬 방식에 의한 정렬이 이루어져 정렬 망 출력포트에 그림 2와 같이 출력된다.

입력되는 셀 들은 3비트의 목적지 주소를 가지고 입력되며, 출력포트에서는 1비트의 목적지 주소가 추가된 4비트 목적지 주소(상단 = 0, 하단 = 1추가)로 바뀌어 출력된다.

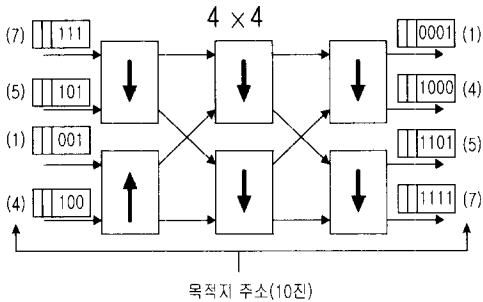


그림 4. 4x4 정렬 망의 구조

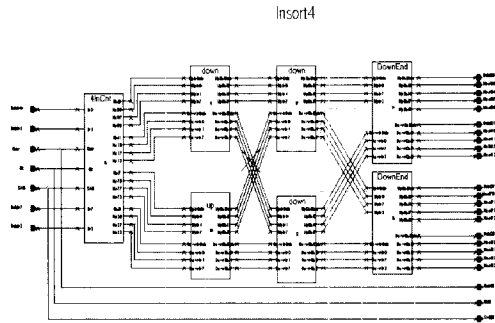


그림 5. 4x4 정렬 망 회로 설계

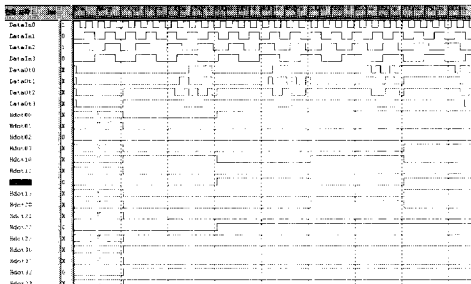


그림 6. 4x4 정렬 망 시뮬레이션 결과

2) 스위치 망의 동작 및 설계

스위치 망은 그림 3 과 같이 기존의 반안 망에 한 개의 단을 추가하여 망을 구성하고 자기경로 배정방식으로 동작시킨다.

망의 처음 부분에는 4x4 교환 스위치 2단으로 구성되어 있으며, 이 스위치들은 반안 망 전단에 추가된 4x4 정렬 망에서 입력된 출력주소를 가지고 자기경로 배정방식에 따라 경로를 배정한다.

이 때 입력 단의 셀 들이 정렬되어 있으므로 역반안 망에 경로 배정하여 입력 셀 들이 충돌하는 것을 방지한다.

이렇게 4x4를 통과한 셀 들은 분배 스위치에서 상, 하 스위치로 분배되는데 상단의 4x4 스위치(1)에서 전송된 셀 헤더(목적지 주소)는 2번째 비트에 0 이 추가된 상태이고 하단의 4x4 스위치(2)에는 전송된 셀 헤더(목적지 주소)의 2번째 비트에 1이 추가된 상태이므로 충돌 없이 4번째 스위치로 전송한다.

4번째 단의 스위치는 출력 버퍼를 내장하고 있어 입력이 2개일 경우 입력되는 셀 들을 처리하게 된다. 스위치 망은 앞에서 소개한 바와 같이 기존 8x8 반안 망에 스위치 소자 1단을 추가하여 반안 망을 병렬로 연결한 것과 같이 동시에 두 개의 경로를 갖도록 설계하였다.

제안된 반안 망은 목적지 주소가 같은 셀 들이 입력되었을 경우의 그림 3과 같이 경로 배정함으로써 두 개의 경로를 갖는다.

이는 결과적으로 8x8 반안 망을 두 개를 연결하여 사용하는 효과를 가진다.

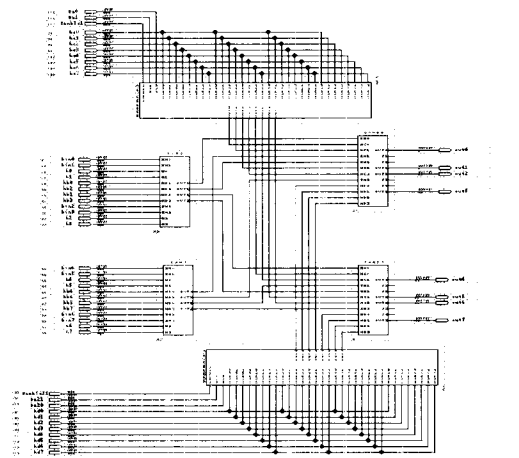


그림 7. 8x8 스위치 망 회로 설계

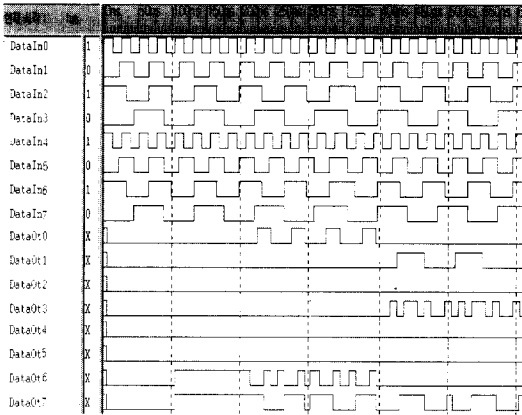


그림 8. 8x8 반얀 망 시뮬레이션 결과

그림 8은 입력된 데이터는 상위 블록과 하위블록에 같은 값을 입력하여 2개의 같은 목적지 주소 가진 데이터가 셀 손실 없이 경로 배정되어 출력되는 것을 볼 수 있다.

IV. 스위치의 성능분석

앞에서 소개한 바와 같이 같은 목적지를 가진 두 개 이상의 셀들이 반얀 망을 통과하려면 내부충돌로 인하여 성능이 저하되기 때문에 성능을 평가해본 후 스위치를 설계할 필요성이 있다. 본 장에서는 스위치 망의 성능과 하드웨어 양을 비교 분석하고 하드웨어의 논리시뮬레이션 결과를 설명한다.

4.1 셀 처리율

입력 버퍼형 스위치 망의 셀 처리율을 살펴보면 반얀 형 스위치 망의 성능을 평가할 수 있다^{[10][11]}. 입력 버퍼형 방식은 입력 단에 버퍼를 두는 방법으로 스위칭 시스템 내의 연결 경로가 함당되지 않거나 일시적으로 차단되었을 경우 셀들을 저장하는데 이용하게 된다. 이러한 시스템은 HOL(Head Of Line) 충돌에 의해 통과하거나, 머무르게되는 이항 확률적 상황을 갖는다. 따라서 셀 처리율을 다음 식으로 표현한다.

$$Pr[A = k] = \left(\frac{N}{k}\right) \left(\frac{p}{N}\right)^k \left(1 - \frac{p}{N}\right)^{N-k} \quad (2)$$

단 $k=0, 1, 2, \dots, N$ 이다.

(N : 입력 수 k :스위치 내에 입력되는 셀 수, p : 유효부하)

반일 입력포트에 도착한 셀들 중 목적지가 같은 두 개 이상의 셀들을 동시에 처리할 수 있다면 식

(3)과 같이 표현할 수 있다.

유효부하를 p 라 할 때, 교환 시스템에 입력되는 셀 수가 k 일 확률 $Pr[A=k]$ 는

$$Pr[A = k] = \left(\frac{N/b}{k}\right) \left(\frac{p}{N/b}\right)^k \left(1 - \frac{p}{N/b}\right)^{N/b-k} \quad (3)$$

단 $k=0, 1, 2, \dots, 2N$ 이다.

식 (3)에서 b 는 동시에 처리될 수 있는 셀들의 수이고 동일한 출력포트의 주소를 갖는 셀들이 한 타임 슬롯에 도착하는 경우의 셀 손실은 $k-b$ 개이다.

따라서 포트수가 N 인 입력 버퍼형 교환 시스템의 셀 손실 확률 $Pr[Cell\ loss]$ 는

$$Pr[Cellloss] \quad (4)$$

$$= \frac{1}{p} \sum_{k=b}^{N/b} (k-1) \left(\frac{N/b}{k}\right) \left(\frac{p}{N}\right)^k \left(1 - \frac{p}{N}\right)^{N/b-k}$$

$$= 1 - \frac{1}{p} \sum_{k=b}^{N/b} (k-1) \left(\frac{N/b}{k}\right) \left(\frac{p}{N}\right)^k \left(1 - \frac{p}{N}\right)^{N/b-k}$$

(4)는 제안된 정렬 반얀 망의 셀 손실 공식 이고 $N \rightarrow \infty$ 일 때

$$Pr[Cellloss] = 1 - \frac{1}{p} + \frac{e^{-bp}}{bp} \sum_{k=0}^{b-1} (b-k) \frac{(bp)^k}{k!} \quad (5)$$

이다. 동시 처리되는 셀 수에 따른 스위치 망의 최대 처리율은 표1과 같다. 표에서 보는바와 같이 8x8 스위치 망의 최대처리율은 같은 주소를 가진 셀들이 한 개만 통과할 때는 최대처리율이 62%인 반면, 본 논문에서 제안한 구조를 가진 스위치 망은 두 개의 셀들을 동시에 처리할 수 있으므로 최대 처리율이 66%까지 향상되었음을 확인하였다.

표 6. 동시에 처리되는 셀 수(b) 최대처리율

b \ N	1	2	4	8
8	0.62	0.66	0.75	1.00
16	0.60	0.62	0.66	0.75
32	0.59	0.60	0.62	0.66
64	0.59	0.59	0.60	0.62
128	0.59	0.59	0.59	0.60

4.2 스위치 망의 하드웨어 복잡도 비교

각 스위치별 하드웨어 복잡도를 비교해 보면 N 이 입출력 포트 수일 때 크로스바 스위치 망 $N \times N$, 반얀 스위치망 $N \log_2 N$, 배저 - 반얀스위치망 $(\log_2 N \cdot (\log_2 N + 1)) / 2 + \log_2 N$, 제안된 정렬스위치 망 $(\log_2 N \cdot (\log_2 N - 1)) / 2 + 2 \log_2 N$ 로 구성되며 표 2와 같다.

